基于互补约束规划模型的投影 Barzilai-Borwein 梯度算法求解绝对值方程 *

严 涛

(南京理工大学 理学院 数学系，南京 210014)

摘 要：在绝对值方程 $Ax-|x|=b$ 问题有解情形下，给出了求解绝对值问题的一新方法。首先建立了一等价求解绝对值问题的互补约束规划模型，进而利用新模型中的非负约束，给出了求解绝对值问题投影 Barzilai-Borwein(BB)梯度算法。数值实验结果表明了方法的有效性。

关键词：绝对值方程；互补约束规划；投影梯度

中图分类号：TP 301.4; O242.2　　doi: 10.19734/j.issn.1001-3695.2018.08.0765

Projected Barzilai-Borwein gradient method for absolute value equations based on mathematical programs with complementarity constraints model

Yan Tao

(*Dept. of Mathematics*, *College of Science*, *Nanjing University of Science & Technology*, *Nanjing* 210094, *China*)

Abstract: This paper presented a mathematical program with complementarity constraints (MPCC) model to solve the absolute value equation $Ax-|x|=b$. Based on the new model, the projected Barzilai-Borwein (BB) gradient method was applied to solve a penalty subproblem and the solution of the original problem was obtained. Preliminary numerical results illustrated the efficiency of the new method.

Key words: absolute value equation; mathematical program with complementarity constraints; projected gradient method

0 引言

考虑如下绝对值方程问题:

$$Ax-|x|=b \tag{1}$$

其中： $A\in R^{n\times n},b\in R^n$ ， $|x|$ 按 x 分量的绝对值获取。绝对值问题为广义绝对值问题 $Ax+B|x|=b$ 的特殊情形[1]，同时在一定条件下，问题等价于线性互补问题[2,3]。绝对值方程问题已吸引许多学者对其理论和算法进行深入的研究，且已有许多相关成果出现。文献[4]对相关理论进行了详细介绍。基于非光滑方程，文献[5,6]提出了相应的半光滑牛顿法；应用各类光滑函数将原问题转换为光滑方程，文献[7~9]提出了相应的光滑化方法；文献[10]在 A 为对称矩阵情形下，将绝对值方程问题转换为了无约束优化问题，进而提出了求解原问题的 Gauss-Seidel 算法。另一方面，互补约束数学规划是一类约束中含有互补问题的优化问题。它是均衡约束规划的一个特殊子类，且在工程设计、经济建模等实际问题中有着很好的应用[11]。近年来，对于互补约束规划的算法已有很好的研究进展，如罚函数逼近法[12],光滑逼近法[13],正则方法[14]。受上述结果启发，本文建立了求解绝对值方程问题的互补约束规划等价模型，进而构造了罚函数子问题，并用投影 BB 梯度方法求解子问题以期获得原绝对值方程问题的解。

1 互补约束规划模型

首先给出一个与绝对值方程问题等价的约束方程系统，即寻找 $w\in R^{2n}$ 使得

$$\begin{aligned}&\bar{A}w=b,\\&w_i\ge 0,\ \ w_iw_{i+n}=0,i=1,2,\cdots,n,\end{aligned} \tag{2}$$

其中 $\bar{A}=(A-I,-A-I)$ 。

定理 1 如果 x^* 为绝对值方程问题(1)的解，则 $w=\left(\left(\frac{x^*+|x^*|}{2}\right)^T,\ \left(\frac{x^*+|x^*|}{2}\right)^T\right)^T$ 为问题(2)的解；反之，如果 w^* 为问题(2)的解，定义 $x_i=w_i^*-w_{i+n}^*\quad(i=1,2,\cdots,n)$,则 x 为绝对值方程问题(1)的解。

证明 将 w 直接代入问题(2)则得到第一部分结论。对于证明第二部分，首先将 w^* 分解为 $(W_1^{*T},W_2^{*T})^T$ ，其中 $W_1^*=(w_1^*,\cdots,w_n^*)^T,W_2^*=(w_{n+1}^*,\cdots,w_{2n}^*)^T$ ，则问题(2)中的第一个等式可写成 $A(W_1^*-W_2^*)-(W_1^*+W_2^*)=b$ 。 因此，可构造 $x=W_1^*-W_2^*$， 即 $x_i=w_i^*-w_{i+n}^*(i=1,2,\cdots,n)$. 下面只需证 $|x|=W_1^*+W_2^*$ 。由问题(2)有 $w_i^*\ge 0,w_i^*w_{i+n}^*=0$ ， 则

$$\begin{aligned}|x_i|&=\sqrt{(w_i^*-w_{i+n}^*)^2}=\sqrt{w_i^{*2}+w_{i+n}^{*2}-2w_i^*w_{i+n}^*}\\&=\sqrt{w_i^{*2}+w_{i+n}^{*2}+2w_i^*w_{i+n}^*}\\&=\sqrt{(w_i^*+w_{i+n}^*)^2}=w_i^*+w_{i+n}^*\end{aligned}$$

则 $|x|=W_1^*+W_2^*$， 第二部分结论得证。

进一步，本文将问题(2)改写为如下互补约束规划问题：

$$\begin{aligned}\min\quad&\frac{1}{2}\|\bar{A}w-b\|^2\\ s.t.\quad&G(w)\ge 0,H(w)\ge 0,\\&G(w)^TH(w)=0,\end{aligned} \tag{3}$$

其中： $G_i(w)=w_i,H_i(w)=w_{n+i}(i=1,2,\cdots,n)$ 。且 w 为问题(3)的最优解解当且仅当 w 为问题(2)的解。因此，求解绝对值方程问题(1)可转换为求解互补约束规划问题(3)。

收稿日期：2018-08-27；**修回日期：**2018-11-16　**基金项目：**国家自然科学基金资助项目（11671205）

作者简介：严涛（1977-），男，江苏泰兴人，副教授，博士研究生，主要研究方向为最优化理论与算法（tyan@njust.edu.cn）.

2　算法

算法包含两个迭代过程，一是用来更新变量 $w^{(k)}$ 的外迭代，另一个是用来求解每步中子问题的内迭代。为实现求解算法，由文献[15],首先构造如下罚函数问题

$$\begin{aligned} \min\ & \frac{1}{2}\|\bar{A}w-b\|^2+\rho G(w)^T H(w) \\ s.t.\ & G(w)\geq 0, H(w)\geq 0. \end{aligned} \tag{4}$$

其中：ρ 为罚参量。 模型(4)具有互补约束规划问题(3)中的简单非负约束，且可方便应用投影算法，同时本文可以证明通过求解罚函数问题(4)可获得互补约束规划问题(3)的稳定点。因此，算法构造将以模型(4)为基础。

下面给出外迭代算法过程及相应收敛结果。

1）外迭代算法

a)设置初始点 $w^{(1)}=(w_1^{(1)},w_2^{(1)},\cdots,w_{2n}^{(1)})^T\geq 0$，$\rho_1,\beta>1$，令 $k=1$；

b)令 $x_i=w_i^{(k)}-w_{i+n}^{(k)}, i=1,2,\cdots,n$，如果 $\|Ax-|x|-b\|=0$，停止计算;

c)令 $\rho=\rho_k$，以 $w^{(k)}$ 为初始点，求解问题(4)的稳定点 $w^{(k+1)}$；

d)令 $k=k+1,\rho_{k+1}=\beta\rho_k$，转步骤 b)。

定理 2　对于每个 ρ_k，$w^{(k)}$ 为问题(4)的一个稳定点。令 $\rho_k\to\infty$，且假设 $\bar{w}$ 为 $\{w^{(k)}\}$ 的极限点，则 $\bar{w}$ 为均衡约束规划问题(3)的 C-稳定点。

证明　由 $G(w)$ 和 $H(w)$ 的定义可知，$\nabla G_i(\bar{w})(\nabla H_j(\bar{w}))$ 为第 $i(j)$ 个元素为 1，其他位置元素为 0 的向量。则

$$\{\nabla G_i(\bar{w})\,(i\in\mathcal{I}_G(\bar{w})),\nabla H_j(\bar{w})\,(j\in\mathcal{I}_H(\bar{w})),$$

$\nabla G_q(\bar{w}),\nabla H_q(\bar{w})\,(q\in\{q:G_q(\bar{w})>0,H_q(\bar{w})>0\})\}$ 线性独立。由文献[12]的引理 3.2 得，$\bar{w}$ 为互补约束规划问题(3)的可行点。此外，又因为互补约束规划线性独立约束品性在 $\bar{w}$ 成立，即 $\{\nabla G_i(\bar{w}),\nabla H_j(\bar{w}):i\in\mathcal{I}_G(\bar{w}),j\in\mathcal{I}_H(\bar{w})\}$ 线性独立，则由文献[12]的定理 2.1 可知，$\bar{w}$ 为均衡约束规划问题(3)的 C-稳定点。

在外迭代算法中，本文每进行一次迭代都需要求解步骤 c）中的问题(4)。下面给出内迭代算法过程。令

$$f(w)=\frac{1}{2}\|\bar{A}w-b\|^2+\rho G(w)^T H(w)$$

及由 $G(w)$ 和 $H(w)$ 的定义，罚函数问题(4)可简化为如下表达形式：

$$\begin{aligned} \min\ & f(w) \\ s.t.\ & w\geq 0. \end{aligned} \tag{5}$$

在内迭代算法中,本文用投影 Barzilai-Borwein 梯度法[16]求解问题(5)。搜索方向设定为

$$d^{(k)}=P(w^{(k)}-\alpha_{BB}\nabla f(w^{(k)}))-w^{(k)}\quad ,$$

其中: $P(\cdot)$ 为投影算子，α_{BB} 为由 BB 方法获得的步长。此外，为确保全局收敛性，算法沿上述投影方向还采用了非单调线搜索获得线搜索步长 λ，则迭代公式为 $w^{(k+1)}=w^{(k)}+\lambda d^{(k)}$。

为区别内外迭代算法中的变量记号，本文令内迭代的变量为 z，内迭代算法如下：

2)内迭代算法

a)设置初始点 $z^{(1)}=w^{(k)}$，$\alpha_{BB}^{(1)}=1/\|\nabla f(z^{(1)})\|_\infty$，令 $l=1$；

b)如果 $\|P(z^{(l)}-\nabla f(z^{(l)}))-z^{(l)}\|=0$，停止计算，输出 $w^{(k+1)}=z^{(l)}$；否则，令 $d^{(l)}=P(z^{(l)}-\alpha_{BB}^{(l)}\nabla f(z^{(l)}))-z^{(l)}$，$\lambda=1$,转步骤 c);

c)如果

$$\begin{aligned} & f(z^{(l)}+\lambda d^{(l)}) \\ & \quad\leq \max_{0\leq j\leq\min(l,M-1)} f(z^{(l-j)})+10^{-4}\lambda(d^{(l)})^T\nabla f(z^{(l)}), \end{aligned}$$

则令 $z^{(l+1)}=z^{(l)}+\lambda d^{(l)}$，转步骤 d)；否则令 $\lambda=\lambda/4$，重复步骤 c);

d)计算 $s^{(l)}=z^{(l+1)}-z^{(l)}$，$y^{(l)}=\nabla f(z^{(l+1)})-\nabla f(z^{(l)})$。如果 $\langle s^{(l)},y^{(l)}\rangle\leq 0$，令 $\alpha_{BB}^{(l+1)}=10^{20}$；否则令

$$\alpha_{BB}^{(l+1)}=\min\left\{10^{20},\max\left\{10^{-20},\frac{\langle s^{(l)},s^{(l)}\rangle}{\langle s^{(l)},y^{(l)}\rangle}\right\}\right\}.$$

e)令 $l=l+1$，转步骤 b）。

由文献[16]中定理 2.2 有如下结果

定理 3　由内迭代算法产生的序列 $\{z^{(l)}\}$ 的任何聚点都是罚函数问题(4)的稳定点。

3　数值结果

外迭代算法中的停机准则设定为 $\|Ax-|x|-b\|\leq 10^{-6}$。

下面讨论内迭代算法的停机准则。　对于简化的罚函数模型(5)，本文定义 $f(w)$ 投影梯度为

$$(\nabla^p f(w))_i=\begin{cases} (\nabla f(w))_i & w_i>0 \\ min\{0,(\nabla f(w))_i\} & w_i=0. \end{cases}$$

则 w 为问题(5)的稳定点当且仅当 $\|\nabla^p f(w)\|=0$。另一方面由文献[16]的引理 2.1 可知，上面条件等价于 $\|P((w)-\nabla f(w))-w)\|=0$，且文献[17]的数值实验表明使用前者的效果更好，因此，内迭代算法停机准则设定为

$$\|\nabla^p f(w^{(k)})\|\leq\bar{\epsilon}=10^{(-3)}\|\nabla^p f(w^{(1)})\|,$$

其中：$w^{(1)}$ 为外迭代算法中的初始点。同时，如果内迭代算法迭代一次结束，则减少 $\bar{\epsilon}$ 的数值，设定为 $\bar{\epsilon}=\frac{1}{10}\bar{\epsilon}$。同时设定 $\rho_1=1,\beta=1.5$。内迭代中 M=10。设定最大外迭代步数为 50，最大内迭代步数为 200。

本文首先采用 100 个随机产生的绝对值方程问题作为算法测试问题。 为确保问题有解，即使得 A 的奇异值大于 1，本文取 A=$U\Sigma V$，其中 U 和 V 为均匀分布在[-10,10]的正交矩阵，对角阵 Σ 的元素从区间[1,50]中随机选取。 取均匀分布在[-1,1]的向量 x^*，则 $b=Ax^*-|x^*|$。外迭代算法的初始值 $w^{(1)}=$|rand(2n,1)-rand(2n,1)|，本文算法在达到设定的误差情形下成功求解了所有测试题，具体实验结果如表 1 所示。

表 1　100 个随机测试题计算结果

Table 1　Results of 100 random test problems

测试题	结果
测试题维数 n	1000
算法成功求解个数	100
$\|Ax-\|x\|-b\|$ 平均值	3.6282e-007
外迭代步数平均值	32.04
$\|x-x^*\|$ 平均值	3.5772e-007

同时，本文也采用了如下两个算例进行数值验证。

例 2[18]　矩阵 A 定义为 $a_{ii}=4n,a_{i,i+1}=a_{i+1,i}=n,a_{ij}=0.5,(i,j=1,2\cdots n)$，$b=(A-I)e$，其中 I 为 n 维单位矩阵，e 为元素全为 1 的列向量。该问题的唯一解为 $x=(1,1,\cdots 1)^T$。初始点设定为零向量，算法对 $n=200,n=500$, 及 n=1000 的维数问题进行测试且获得了问题的唯一解，其他迭代次数、停机时的罚参量数值及 $\|Ax-|x|-b\|$ 数值如表 2 所示。

表 2　例 2 计算结果

Table 2　Results of example 2

维数 n	迭代步数	罚参量 ρ	$\|Ax-\|x\|-b\|$
200	22	194.6195	2.7181E-007
500	29	2.2168E+003	1.8910E-007
1000	29	3.3253E+003	6.2445E-007

例 3[19]　矩阵 A 定义为 $a_{ii}=500,a_{ij}=a_{ji}=1+rand,(i\neq j)$。其中

rand 为[0,1]的随机数。令 $b=(A-I)e$，其中 I 为n维单位矩阵，e 为元素全为 1 的列向量。该问题的唯一解为 $x=(1,1,\cdots 1)^T$ 。本文算法对 $n=200, n=500,$ 及 $n=1000$ 维数问题进行测试。初始点设为零向量。经过有限次迭代，算法获得问题的唯一解，迭代次数、停机时的罚参量数值及 $\|Ax-|x|-b\|$ 数值如表 3 所示。

表 3 例 3 计算结果

Table 3 Results of Example 3

维数 n	迭代步数	罚参量 ρ	$\|Ax-\|x\|-b\|$
200	20	129.7463	5.0775E-007
500	25	656.8408	1.8032E-007
1000	23	194.6195	4.7039E-008

4 结束语

本文给出了求解绝对值方程问题基于均衡约束规划的投影梯度型算法。数值实验验证了算法的有效性，也表明本文算法为求解绝对值方程问题的一个可行方法。

参考文献：

[1] Rohn, J. A theorem of the alternatives for the equation Ax+B|x|=b [J]. Linear and Multilinear Algebra, 2004, 52 (6): 421-426.

[2] Mangasarian O L. Absolute value programming [J]. Computational Optimization and Application, 2007, 36 (1): 43-53.

[3] Mangasarian O L. Absolute value equation solution via concave minimization [J]. Optimization Letters, 2007, 1 (1): 3-8.

[4] Mangasarian O L, Meyer R R. Absolute value equations [J]. Linear Algebra and its Application, 2006, 419 (5): 359-367.

[5] Mangasarian O L. A generalized Newton method for absolute value equations [J]. Optimization Letters, 2009, 3 (1): 101-108.

[6] Zhang Chao, Wei Qingju. Global and finite convergence of a generalized Newton method for absolute value equations [J]. Journal of Optimization Theory and Applications, 2009, 143 (2): 391-403.

[7] Louis C, Qu Biao, Zhou Guanglu. A lobally and quadratically convergent method for absolute value equations [J]. Computation Optimization Apllication, 2011, 48 (1): 45-58.

[8] Yong Longquan, Liu Sangyang, Zhang Shemin. Smoothing Newton method for solving absolute value equations [J]. International Journal of the Physical Science, 2011, 6 (23): 5399-5405.

[9] 邓永坤，王海军，张萍. 基于极大熵牛顿法求解绝对值方程 [J]. 计算机应用研究，2012，29 (12)：4546-4548. (Deng Yongkun, Wang Haijun, Zhang Pimg. Maximum entropy Newton method for absolute value equations [J]. Application Research of Computers, 2012, 29 (12): 4546-4548.)

[10] Muhammad Aslam Noor, Javed Iqbal, Khalida Inayat Noor, *et al.* On an iterative method for solving absolute value equations [J]. Optimization Letters, 2012, 6 (5): 1027-1033.

[11] Luo Zhiquan, Pang Jongshi, Ralph D. Mathematical programs with equilibrium constraints [M]. Cambridge:: Cambridge University Press, 1996.

[12] Hu Xinmin, Ralph D. Convergence of a penalty method for mathematical programming with complementarity constraints [J]. Journal of Optimization Theory and Applications, 2004, 123 (2): 365-390.

[13] Facchiner F, Jiang Houyuan, Qi Liqun. A smoothing method for mathematical programs with equilibrium constrains [J]. Mathematical Programming, 1999, 85 (1): 107-134.

[14] Scholtes S. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints [J]. SIAM Journal on Optimization, 2001, 11 (4): 918-936.

[15] Tin-Loi, F. On the numerical solution of a class of unilateral contact structural optimization problems [J]. Structural Optimization, 1999, 17 (2-3): 155-161.

[16] Birgin E G, Martinez J M, Raydan M. Nonmonotone spectral projected gradient methods on convex sets [J] . SIAM Journal on Optimization, 2000, 10 (4): 1196-1211.

[17] Lin zhiren, More J J. Newton's method for large bound-constrained optimization problems [J] .SIAM Journal on optimization, 1999, 9 (4): 1100-1127.

[18] Ujevic N. A new iterative method for solving linear systems [J]. Applied Mathematics and Computation, 2006, 179 (2): 725 –730.

[19] Burden R L, Faires J D. Numerical analysis [M]. 7th ed. Boston: PWS Publishing Company, 2006.